Army Materiel Systems Analysis Activity

AMSAA TECHNICAL REPORT NO. TR-2012-65

SCHEDULE RISK DATA DECISION METHODOLOGY (SRDDM)

September 2012

**U.S. ARMY MATERIEL SYSTEMS ANALYSIS ACTIVITY
ABERDEEN PROVING GROUND, MARYLAND 21005-5071**

DESTRUCTION NOTICE

Destroy by any method that will prevent disclosure of contents or reconstruction of the document.

DISCLAIMER

The findings in this report are not to be construed as an official Department of the Army position unless so specified by other official documentation.

WARNING

Information and data contained in this document are based on the input available at the time of preparation.

TRADE NAMES

The use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial hardware or software. The report may not be cited for purposes of advertisement.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (LEAVE BLANK)	2. REPORT DATE September 2012	3. REPORT TYPE AND DATES COVERED Technical Report	
4. TITLE AND SUBTITLE Schedule Risk Data Decision Methodology (SRDDM)			5. FUNDING NUMBERS
6. AUTHOR(S) John Nierwinski, Jr.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Director, U.S. Army Materiel Systems Analysis Activity 392 Hopkins Road Aberdeen Proving Ground, MD 21005-5071			8. PERFORMING ORGANIZATION REPORT NUMBER TR-2012-65
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED			12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

One of the top priorities of the U.S. Army is to make decisions regarding acquisition programs that will best serve the Warfighter. Providing an accurate and precise schedule risk assessment for a set of alternatives is a key input to the decision making process. Weapon System Acquisition Reform Act of 2009 is driving more analysis to support the Analysis of Alternative (AoA).

AMSAA conducts independent schedule risk assessments to support AoAs and other major Army acquisition studies. A probability is assessed for completing a given phase within the schedule developed by the Program Manager (PM). The probabilities are based upon historical data for analogous programs.

AMSAA developed a Schedule Risk Data Decision Methodology (SRDDM) that determines if enough historical data exists to utilize quantitative techniques to conduct the schedule risk assessment. This methodology lays the mathematical and decision-making foundation for all future schedule risk assessments. SRDDM uses Monte Carlo simulations and mathematical models to build a confidence interval (CI) for the probability of meeting the PM's schedule. If this CI width is within tolerance then enough analogous programs exist to build risk distributions.

AMSAA has applied SRDDM to the Indirect Fire Protection Capability and the Armored Multi-Purpose Vehicle AoAs. Future work includes risk mitigation, trade space analysis, and developing event-driven models.

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAME AS REPORT

THIS PAGE INTENTIONALLY LEFT BLANK.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

LIST OF ACRONYMS

AMPV	- Armored Multi-Purpose Vehicle
AMSAA	- Army Materiel Systems Analysis Activity
AoA	- Analysis of Alternatives
BC	- Bias Corrected
CI	- Confidence Interval
DM	- Decision Maker
FUE	- First Unit Equipped
IFPC	- Indirect Fire Protection Capability
IOC	- Initial Operational Capability
KS	- Kolmogorov-Smirnov Goodness of Fit
LB	- Lower bound
P	- Probability of meeting PM's schedule
PM	- Program Manager
SME	- Subject Matter Expert
SPSD	- Schedule Proportion Sampling Distribution
SRDDM	- Schedule Risk Data Decision Methodology
UB	- Upper bound

LIST OF SYMBOLS

$\hat{P}$	Probability estimate of meeting PM's schedule
α_s	Significance level
$\hat{z}_0$	The value of Bias Correction
$\Phi(\ )$	Standard normal cumulative distribution function
$\Phi^{-1}(\ )$	Inverse of $\Phi(\ )$
$\hat{P}^{(\alpha_s)}$	$100*\alpha_s$ th percentile from the 500 estimates of P
$\hat{P}^{(\alpha_2)}$	$100*\alpha_2$ th percentile from the 500 estimates of P
$z^{(1-\alpha_s)}$	$100*(1-\alpha_s)$th percentile point of a standard normal distribution
$\hat{P}_i$	Probability estimated for the ith sample.
O(1/n)	Second order accuracy
$O(\frac{1}{\sqrt{n}})$	First order accuracy
Z	From the standard normal distribution with area $\alpha_s/2$ in the tail.
n	Number of analogous programs
σ^2	$\hat{P}*(1-\hat{P})$

SCHEDULE RISK DATA DECISION METHODOLOGY (SRDDM)

1.   INTRODUCTION

One of the top priorities of the U.S. Army is to make decisions regarding acquisition programs that will best serve the Warfighter. Providing an accurate and precise schedule risk assessment for a set of alternatives is a key input to the decision making process. The Weapon System Acquisition Reform Act of 2009 requires trade-offs among technical, schedule, & cost risks along with performance to support the Analysis of Alternative (AoA). This requirement remains critical for all defense acquisition programs [1].

AMSAA conducts independent schedule risk assessments to support AoAs and other major Army acquisition studies. A probability is assessed for completing a given phase (e.g. Milestone B to C) within the schedule developed by the Program Manager (PM). The probabilities are based upon historical data for analogous acquisition programs.

Analogous acquisition programs are historical programs or elements of historical programs exhibiting characteristics that are relatively similar to a specific AoA alternative. Some of these characteristics include program type, acquisition strategy, system capabilities, critical technologies, and schedule drivers.

AMSAA developed a Schedule Risk Data Decision Methodology (SRDDM) that determines if enough historical data exists to utilize quantitative techniques to conduct the schedule risk assessment. This methodology lays the mathematical and decision-making foundation for all future schedule risk assessments. Within SRDDM are Monte Carlo simulations and mathematical models that build a confidence interval (CI) around the probability of meeting the PM's schedule. If the CI width is within the user established tolerance then enough analogous programs exists to build risk distributions. If these distributions represent the risks for the alternative then the distribution associated with the smallest CI width is chosen.

The main goal or purpose for schedule risk assessments is to accurately and precisely assess the probability of meeting the PM's schedule, where a low probability is an unfavorable outcome or high risk. The importance or meaning of meeting this schedule (i.e. mild to severe consequences) will be determined by Army decision makers.

AMSAA has applied SRDDM to the Indirect Fire Protection Capability (IFPC) and the Armored Multi-Purpose Vehicle (AMPV) AoAs. Risk mitigation and trade space analysis can be performed using SRDDM. Future work includes developing event-driven models and incorporating technical risk assessment outcomes and subject matter expert (SME) input.

2. SRDDM METHODOLOGY

2.1 Schedule Risk Data.

Schedule risk assessments can utilize two levels of historical data:

- Phase-level data: High level (i.e. Milestone phase dates) series relationships between schedule events.
- Event-level data: Detailed (series and or parallel) relationships between schedule events – this includes Work Breakdown Structure and Critical Path analyses.

Currently SRDDM focuses on phase-level approaches because the data is more readily available than event level information.

SRDDM uses two data modeling techniques with the phase-level data resolution to create distributions - Ratio and Unadjusted. The Ratio technique calculates the percent change between estimated and actual dates from historical analogous programs and applies it to the PM's estimate for the alternative. The Unadjusted technique uses actual dates from the historical programs.

The figure below illustrates phase-level Ratio and Unadjusted techniques:
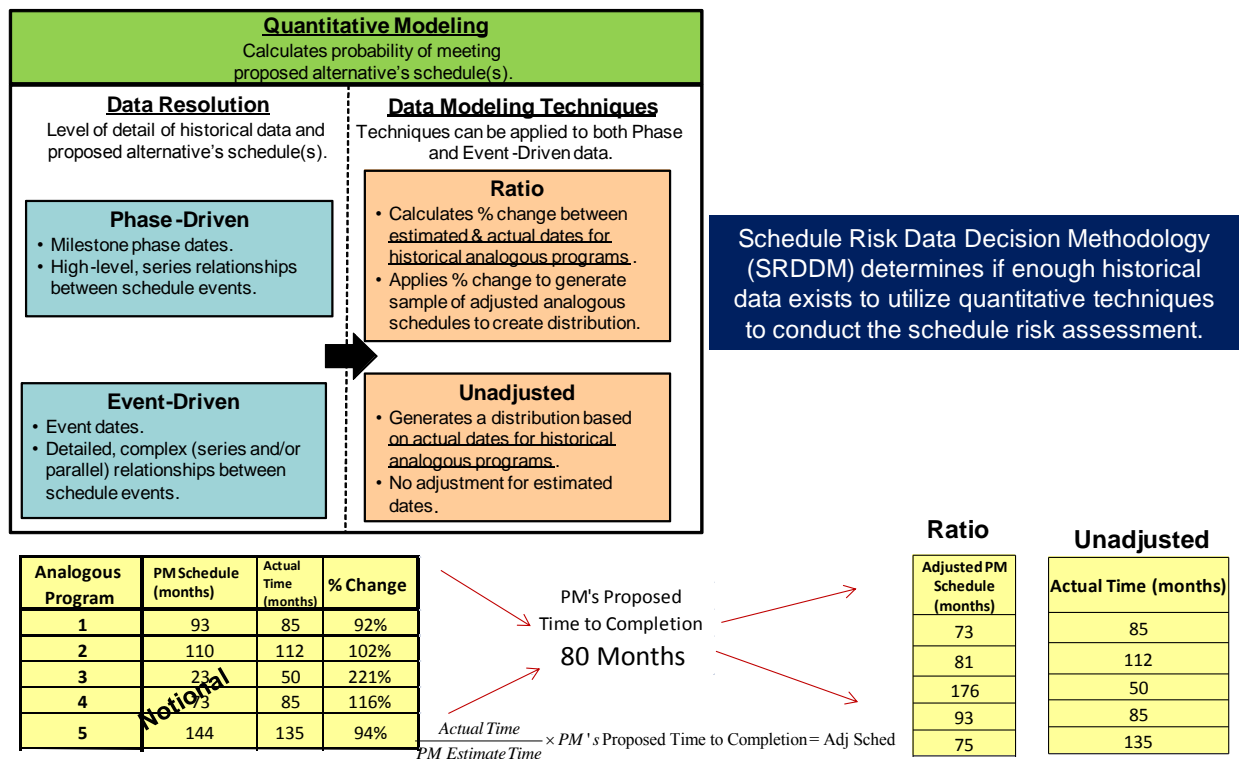

Figure 1. Schedule Risk Data.

2.2 Schedule Risk Distribution.

This document focuses on phase-driven schedule risk distributions. Distributions are created for each of the alternatives. These distributions can be fitted parametrically (e.g. lognormal) using the "best fitting" distribution or non-parametrically (empirical data).

There are two data modeling techniques and the option of "best fitting" distribution or empirical distribution to measure risk. Hence, there are four phase-driven distributions as shown in the tree figure below.
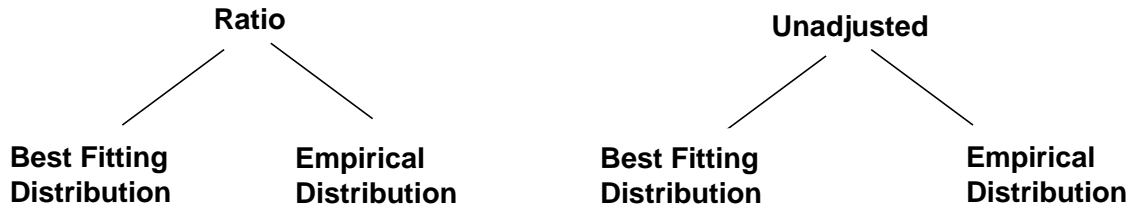
Figure 2. Phase Distribution Tree.

Two classical methods are used to determine if the sample times from analogous programs fit a parametric distribution. First, the Kolmogorov-Smirnov (KS) Goodness of Fit Test is evaluated because the samples of analogous programs are small [2]. If the KS Test results for a given distribution reveals a significant result [i.e. P-value must be less than some significance level (α_s)] then this distribution is considered as a potential candidate for the "best fitting" distribution. For this hypothesis test and application, $\alpha_s =$ 10% is commonly used. Second, the Quantile-Quantile (Q-Q) plot is examined. A Q-Q plot graphs the quantiles of the empirical data against the parametric distribution. Quintiles are data points from the cumulative distribution taken at regular probability intervals. For example, quintiles with 100 points are called percentiles and quintiles with 10 points are called deciles. If the data fits the parametric distribution, then the Q-Q plot should be a straight line [2]. If two or more distributions have a straight line Q-Q plot and P-value < α_s, then choose the distribution with the smallest P-value.

2.3 Sufficiency of Historical Data.

2.3.1 Objective and Background.

Evaluation must be completed to determine if there are enough analogous programs (for a given alternative) to apply the distributions. Error must be strongly considered regarding the assessed probability (let's call it $\hat{P}$) of completing a given phase (i.e. Milestone B to C, etc.) within the schedule developed by the PM.

So, a confidence interval (CI) for P is needed, where P is the true but unknown probability of meeting PM's schedule. $\hat{P}$ and the sample are used to build the CI.

Below is a flowchart figure of a Monte Carlo simulation to obtain this CI for a given alternative and for a given distribution:

I. Collect Analogous Data

II. Build distribution

III. Obtain $\hat{P}$ based
on the Alternative estimate.

IV. Resample from the distribution
(Draw a random sample)

V. Obtain a $\hat{P_1}$ based on this random sample.

VI. Repeat steps IV and V 500+ times and obtain 500
estimates $\hat{P_1}, \hat{P_2}, ..., \hat{P}_{500}$

VII. Using the Percentile Method
with Bias Correction (BC Method)
obtain a CI for P (80%, 90%, etc).

Schedule Proportion
Sampling Distribution (SPSD)

Lower Bound (LB)   Upper Bound (UB)

A Figure 3. Confidence Interval Process.

For a given alternative, this simulation is run for each phase-driven distribution.

2.3.2 Schedule Proportion Sampling Distribution.

The Schedule Proportion Sampling Distribution (SPSD) uses Visual Basic and @Risk to compute a sampling distribution for the probability of meeting the PM's schedule. This algorithm uses Monte Carlo simulation [3], resampling methods such as parametric and non-parametric bootstrapping, KS Goodness of Fit testing, Q-Q plotting, and other mathematical tools. This method produces a large number of estimates for P. At least 500 simulation runs (denoted at 500 +) are required for stable results - Step VI from Figure 3.

The first step in this algorithm is to use the sample of analogous programs (sample size is n) to determine if a distribution can be fitted using the KS Goodness of fit test and Q-Q plotting. If none can be fitted, then we use the empirical data as it is.

A random deviate of size n can be drawn from the fitted distribution (parametric bootstrap) or a nonparametric bootstrap of size n can be generated from the empirical data. A nonparametric bootstrap sample of n programs is a random sample of one program at a time with replacement from the original sample of n programs [4]. Once the sample is drawn we compute $\hat{P_1}$ using the data from the first drawn sample. Then go back to the distribution or empirical data from the first step and draw another random sample using either the parametric or nonparametric bootstrap. Now, $\hat{P_2}$ is computed

4

and this procedure is repeated 500+ times to create the sampling distribution for the proportion.

This general concept of resampling using information from the original sample to create a sampling distribution is used in several other AMSAA developed methodologies [4], [5], [6].

2.3.3 Percentile CI with Bias Correction.

The next step is to apply the Bias Corrected (BC) method [7] to this distribution of 500 + estimates of P. The BC method is basically an adjustment (for a non-normal sampling distribution) to the percentile points of the Percentile Method. This method adjusts these percentile points when the mean and median are not equal – hence the method tries to normalize the distribution. In other words, the CI would shift toward the mean, left or right depending on where the median is located. This shift could be wider when there is more skewness in the sampling distribution. This BC adjustment could improve the coverage properties for the confidence interval (CI) problem.

Let $\hat{P}^{(\alpha_s)}, \hat{P}^{(1-\alpha_s)}$ indicate the $100*\alpha_s$th and $100*(1-\alpha_s)$th percentiles from the 500 estimates of P. This represents the percentile method for a 2-sided $100*(1-2\alpha_s)$ CI. The lower and upper bound using the BC method is given by:

$$\hat{P}^{(\alpha_1)}, \text{ where } \alpha_1 = \Phi(2\hat{z}_0 + z^{(\alpha_s)}) \text{ ; Lower Bound}$$

$$\hat{P}^{(\alpha_2)}, \text{ where } \alpha_2 = \Phi(2\hat{z}_0 + z^{(1-\alpha_s)}) \text{ ; Upper Bound}$$

Here $\Phi(*)$ is the standard normal cumulative distribution function and $z^{(1-\alpha_s)}$ is the $100*(1-\alpha_s)$th percentile point of a standard normal distribution. For example $z^{(.95)} = 1.645$ and $\Phi(1.645) = .95$.

The value of BC is derived by the proportion of replications that is less than the original estimate $\hat{P}$. Here is that value [5]:

$$\hat{z}_0 = \Phi^{-1}(\frac{\#\{\hat{P}_i < \hat{P}\}}{500+})$$

2.3.4 Wilson Score Confidence Interval Method.

When the data shows that the probability of meeting the PM's schedule is either near 0 or near 1 then building a sampling distribution for the probability can be difficult. In these

cases using the Monte Carlo simulation approach as documented in the prior sections may be difficult.

In 1927, Wilson developed a confidence interval method [8] for proportions to account for extreme probabilities. Extreme probabilities are those probabilities that demonstrate an unusual high or low probability of not meeting PM's schedule, such as near 0 or near 1. The formula below represents Wilson's definition of a confidence interval where extreme proportions exist.

$$\frac{\hat{P} + \frac{Z^2}{2n} \pm Z\sqrt{\frac{\sigma^2}{n} + \frac{Z^2}{4n^2}}}{1 + \frac{Z^2}{n}}$$

where:

Z = from the standard normal distribution with area $\alpha/2$ in the tail.

n = number of analogous programs

$\hat{P}$ = Probability estimate from the analogous data.

$\sigma^2 = \hat{P} * (1 - \hat{P})$

The confidence level is $100*(1 - \alpha)\%$.

2.3.5 Coverage Validation & Accuracy.

In order to accurately build this 2-sided CI stochastic model, enough sample data is needed to achieve the requested level of confidence (e.g. 90%). Coverage models are used to validate the model accuracy.

First let's define coverage and accuracy. Coverage is defined to be the percentage of CI's that contain the true population parameter P, where each CI is constructed with some method at the $100*(1 - \alpha_s)$th confidence level for a given random sample of n analogous programs. In other words, we need to run the inference method (Monte Carlo simulation with BC method) 500+ times (500+ samples drawn from a parametric or nonparametric population) to obtain 500+ CI's. These 500+ samples are not to be confused with the 500+ iterations from the Monte Carlo simulation with BC method. Inspection is made to determine how many CI's contain the true P.

The decision maker needs to determine how much absolute relative error for confidence is tolerable. The assessment of this error is very similar to the elicitation process to determine CI width in the next section. Most applications of this nature historically have required either 80% or 90% confidence with a 5% absolute relative error when measuring acceptable coverage.

Accuracy is just a convergence rule for explaining the relative error of a 1-sided coverage. The rule focuses on the speed at which the relative error approaches 0. Second

order accuracy is defined as the actual non-coverage probability intended to be α_s% for a 1-sided $(1-\alpha_s)$% CI, approaches the ideal of α_s% with error proportional to 1/n [7]. First order accuracy would approach the ideal of α_s% with error proportional to $\frac{1}{\sqrt{n}}$. This means that the relative error of the 1-sided coverage is of the order O(1/n) for second order accuracy and O($\frac{1}{\sqrt{n}}$) for first order accuracy. The BC method is second-order accurate since the method adjusts the percentile points based on the non-normal sampling distribution. The percentile method is only first-order accurate since it does not make any adjustments to the percentile points.

Lessons learned from a coverage validation study reveals the following results:

- At least 6 analogous programs (n) are needed to perform any of these confidence interval methods.
- If the probability is extreme (near 0 or 1) then use the Wilson Score Interval.
- If the probability is not extreme then use one of the two Monte Carlo methods:
    - Use Percentile Method if n is 10 or less.
    - Use Bias Correction Method if n > 10.

Lessons learned demonstrated that both empirical and best fitting distribution techniques yielded similar coverages. Hence, choose the smallest CI width when selecting between these two techniques.

2.3.6   Precision Error Tolerance.

The decision maker (DM) must decide an acceptable and tolerable width of the CI. The assessment of the "tolerance of width of the CI" is a decision problem which requires proper consideration of what happens to the "big picture problem" if the endpoints of this CI (namely the UB and LB) are truly realized. In other words, the DM may change the decision as a result of the LB or UB occurring. If the decision is changed, then the sensitivity of this width is too large and cannot be tolerated. Hence, the width needs to be smaller. In order to reduce the width, more analogous data needs to be collected.

On the other hand, if the DM does not change the decision as result of this width then the width is acceptable or tolerated, and enough data was collected. Keep in mind that different problems have different sensitivities to CI width. Sometimes a probability of 90% vs. 70% of meeting schedule will not change the overall alternative level decision (i.e. both are directionally pretty good with low risk). However, a probability of 99% vs. 90% of a bridge breaking in the next year could be a decision changer.

For schedule risk assessment applications, the main concern that the decision maker has is on the LB because that is where the risk is contained. Therefore, the risk is greater when a large width exists between the mean and the LB probabilities compared to the UB. The DM needs to assess the largest width (mean to LB) that he or she can live with.

In other words, when does the length of the width become an issue or when does it cause the DM to re-consider his or her decision.

2.4   The Best Phase Distribution.

Next, the distribution that best represents the risk of the alternative is selected. Only the distributions with an acceptable error for the P CI are considered.

Determine the trustworthiness of the analogous programs. The answer to this question lies in the definition of "analogous" that was stated in the Introduction. The core of the definition was that the alternative and the historical programs had the following similar characteristics: program type, acquisition strategy, system capabilities, critical technologies, and schedule drivers.

Determine the trustworthiness of the initial forecast. The answer lies within two areas. First, decide if the quality (i.e. same general forecasting approaches are used) of the forecast is consistent over all historical programs and the alternative.   Second, find out if the initial forecast for the historical programs were truly initial – meaning that the forecast was recorded as initial when in fact it was really after the program started.

If both are trusted, then choose the distribution associated with the smallest CI length for P.

2.5   SRDDM Process Flowchart.

The flowchart in the figure below is a high-level overview of SRDDM.
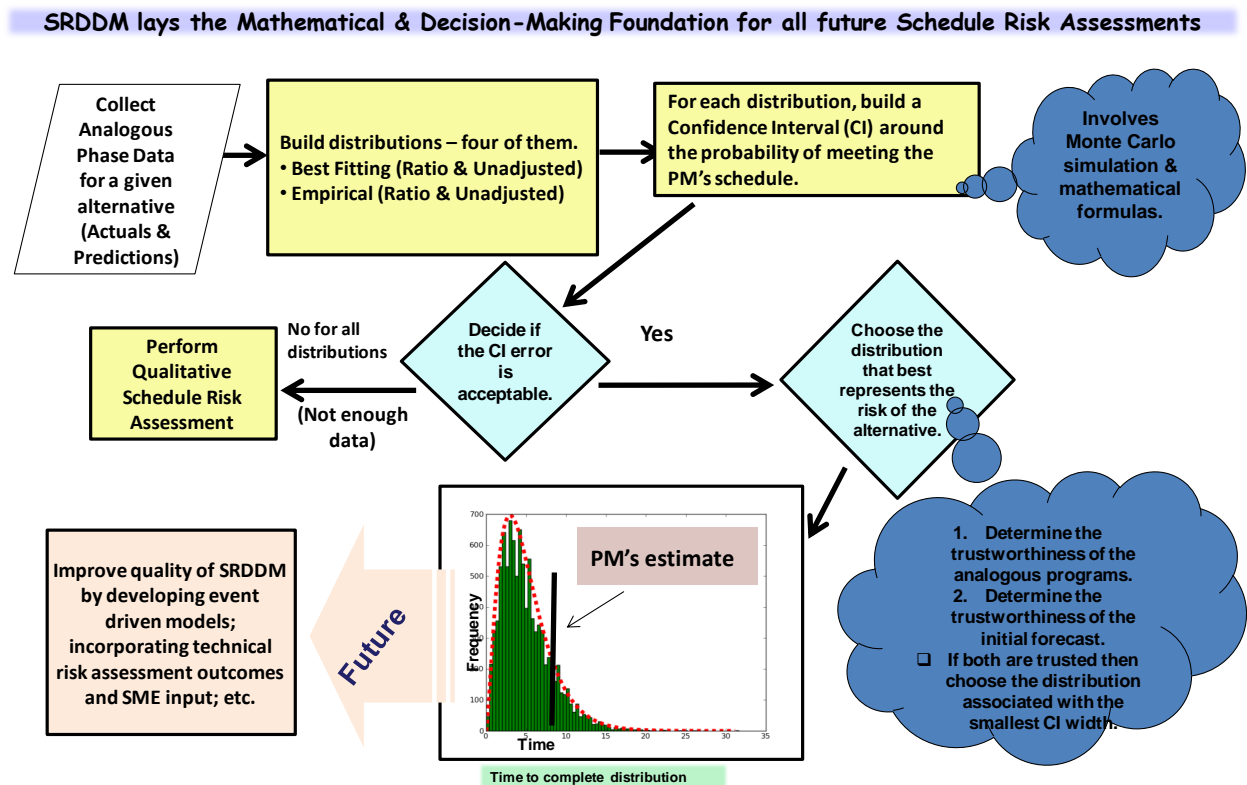
Figure 4. SRDDM Process Flowchart.

The steps to this SRDDM process are:

- Using the analogous data, compute the probability of meeting the PM's schedule ($\hat{P}$). For empirical data, this is the percentage of the analogous data that falls below the PM's schedule. For best fitting distribution, this is the area below the PM's schedule using the best fitting distribution.

- Determine if enough data exists to use $\hat{P}$ as our estimated probability. A CI should be built for meeting the PM's schedule using the analogous data.

- To build this CI we utilize one of three CI methods depending on the number of programs and the suspected probability of meeting the PM's schedule.   These are called:
  - Monte Carlo Simulation Percentile Method
  - Monte Carlo Simulation Bias Corrected Method
  - Wilson Score Interval

  Details of these methods were discussed in earlier sections.

- After using one of these three methods to build a CI, errors (CI width) should be examined. The lower confidence bound (LCB) is of most concern because the LCB represents a higher risk.
- If the data and the initial forecast are trustworthy, then choose the distribution associated with the smallest CI width.
- Risk mitigation and trade space analysis are conducted – discussed in Section 5.

3. NOTIONAL APPLICATION #1

This is a notional example to illustrate the SRDDM process. There are 11 notional analogous programs with this particular alternative within the AoA.

Table 1. Notional Analogous Program Data – Application #1.

Analogous Program	ACTUAL or Unadjusted MS B - IOC (Work Days)	Years	PREDICTED MS B - IOC (Work Days)	Years	Actual/Predicted	Adjusted (Work Days)
1	1787	6.8	2586	9.9	0.69	1866
2	1995	7.6	1763	6.7	1.13	3055
3	1806	6.9	1348	5.1	1.34	3617
4	2437	9.3	2177	8.3	1.12	3022
5	3050	11.6	3785	14.4	0.81	2176
6	1784	6.8	1813	6.9	0.98	2657
7	3218	12.3	3262	12.5	0.99	2664
8	1816	6.9	2002	7.6	0.91	2449
9	2158	8.2	2110	8.1	1.02	2761
10	4435	16.9	5458	20.8	0.81	2194
11	1755	6.7	2044	7.8	0.86	2318
PM estimate =	2700	10.3				

The following results are the outcome when SRDDM is applied to this notional data. The definition and interpretation of these results were explained in detail in Section 2.

Table 2. Application #1 Results.

Distribution	Fit Results	$\hat{P}$	Bias Corrected 80% CI for P	CI Length	Best
Unadj / Fit	None	None	None	None	
Unadj / No Fit	N/A	0.72	(.45 , .82)	0.37	
Ratio / Fit	Gamma	0.63	(.46 , .8)	0.34	Best
Ratio / No Fit	N/A	0.63	(.37 , . 81)	0.44	

There was only one distribution within Unadjusted since the Goodness of Fit test did not reveal a good fit. Note that the CI's for Unadjusted / No Fit and Ratio / Fit are similar only by coincidence. The Bias Correction process worked in opposite directions to cause the Ratio / Fit CI to slide to the right and the Unadjusted. / No Fit CI to slide to the left. This is due to the median value being to the left or right of the mean value, as discussed in Section 2.3.3.

For this notional decision problem, all CI's bounds are within acceptable tolerance and the analogous programs and the quality and consistency of the forecast are trustworthy. Therefore, choose the Ratio / Fit distribution to compute the probability (.63) of meeting the PM's schedule in 10.3 years. Army Senior leadership advises that the uncertainty in our risk assessments be recognized, quantified, and considered. This is exactly what is shown in table 2.

Note that the probability estimates for Ratio and Unadjusted are different due to the fact that the Ratio data includes additional information (i.e. the initial prediction or forecast). This could cause the Ratio probability to be higher or lower than the Unadjusted probability. Everything depends on where the PM's estimate lies. One must not be chosen over another just because it yields a more attractive result. If both distributions are trusted, then choose the distribution associated with the smallest CI width for P. The only reason to override this rule is if one of two things occurs:

1. Trust is low in either the analogous data or the forecast.
2. Trust is much higher in either the analogous data or the forecast.

4.   NOTIONAL APPLICATION #2

Here are results from another notional example which only has actual data and three alternatives.

Table 3. Application #2 Results.

	PM Schedule (in months)		
	MS B to MS C	MS C to FUE	MS B to FUE
	40	**30**	**90**
Alternative A			
# of analogous programs	10	10	8
Probability of Meeting PM Schedule	0.90	0.31	0.60
Confidence Interval (CI)	(0.80, 0.96)	(0.30, 0.62)	(0.43, 0.70)
CI Acceptable ?	Yes	Yes	Yes
Alternative B			
# of analogous programs	13	11	9
Probability of Meeting PM Schedule	0.77	0.20	0.45
Confidence Interval (CI)	(0.68, 0.87)	(0.08, 0.29)	(0.24, 0.58)
CI Acceptable ?	Yes	Yes	Yes
Alternative C			
# of analogous programs	5	5	3
Probability of Meeting PM Schedule	0.60	0	0.70
Confidence Interval (CI)	(0.30, 0.80)	(0, 0.25)	(0.30, 0.90)
CI Acceptable ?	No	No	No
	Not enough data to make a determination		

Notional example based on historical data without initial estimated forecasts; Unadjusted technique used.

Notice that the width of the CI for the 3 or 5 analogous programs used in alternative C is unacceptable because the CI width is too wide and yields too much uncertainty between the probability point estimate and the LCB. Hence, more data is needed. One could always label Alternative C (MS C to FUE) as a high risk (probability = 0) with 5 analogous programs. Directionally, this can only improve with more data. The Wilson Score Interval was used to compute CI's for alternative C since a sampling distribution could not be generated with these few analogous programs. Furthermore, with a probability near 0, the Wilson method would be used anyway.

Recall, coverage lessons learned shows that SRDDM or Wilson method will produce poor coverage validation results with less than 6 analogous programs. Note that the MS B to FUE probability is not always between the MS B to MS C and MS C to FUE

probabilities. Even though the sum of the two proposed schedules for the sub-phases is 90 months, the probability of meeting the schedules of these two sub-phases are sensitive to combinations for obtaining a sum of 90 months.

5.   DATA ALLOCATION ISSUES.

Suppose historical analogous programs from MS B-to-MS C were collected and this data really represented MS A-to-MS C. For example, the programs may have prematurely entered the acquisition process at MS B when the technology readiness levels were actually lower than claimed. This could result in MS A to MS B activities being performed during MS B-to-MS C. An algorithm was designed to allocate some of the time collected in MS B-to-MS C back to MS A-to-MS B. To do this, historical analogous programs are collected that have both phases and weighted average factor is computed to be applied to the time in MS B-to-MS C. This will shift some time back to MS A-to-MS B.

This weighted average factor is based on the history of analogous programs with times in both phases and is only an estimate. Every estimate based on data has a CI associated with it. So, a CI on the factor estimate is computed and then all models are reallocated and re-ran using the mean estimate and the lower and upper bounds from the CI.

Confidence Intervals for Ratio Means (CIM4RM) is used to compute the CI because this metric is a ratio mean [4].

6.   RISK MITIGATION AND TRADE SPACE.

Schedule risk mitigation strategies are used to reduce schedule risk, i.e. make the risk less severe. For example, the DM might want to know what can be done to reduce the schedule risk from high to medium or low risk for a particular alternative within the AoA.

This means that the probability of meeting the PM's schedule ($\hat{P}$) would increase. There are two fundamental ways to solve this risk mitigation problem. Which are:
- Add time to the schedule or
- Remove or reduce schedule drivers – this would cause the "time to complete" distribution (based on analogous programs) to shift to the left.

The following figure illustrates these two risk mitigation approaches.
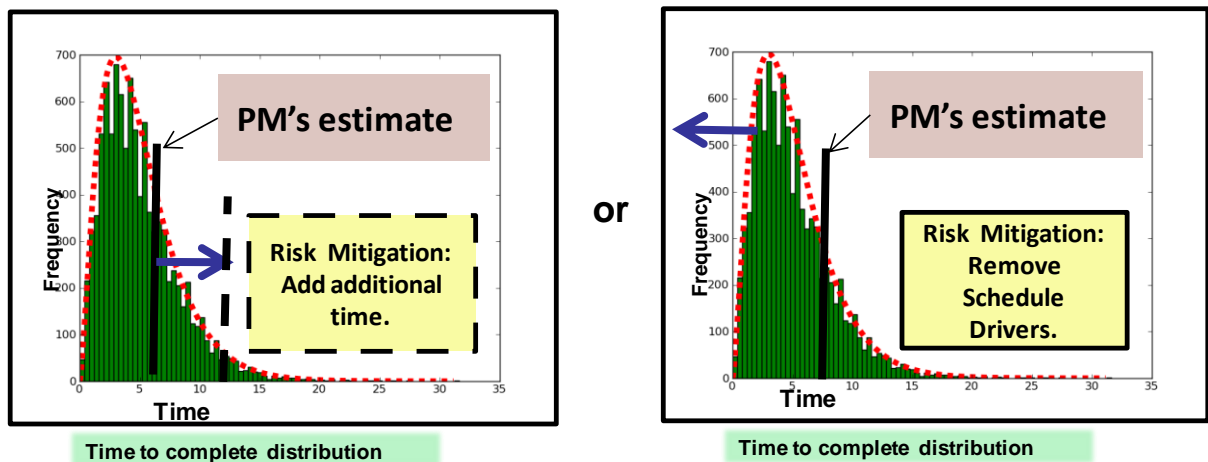
Figure 5. Risk Mitigation Strategies.

Therefore, visualize that $\hat{P}$ would increase in both of these strategies as seen in Figure 5. One can perform sensitivity analyses with both strategies to see how much risk is reduced when time is added to the schedule or schedule drivers are removed or reduced.

However, to successfully achieve risk mitigation within an alternative, one needs to know how much time (if any) should be added to the schedule and how much time (if any) should the distribution shift. In order to achieve this goal, one needs to know the details of the schedule drivers for the analogous programs and for the alternative. These schedule drivers should be similar for both analogous programs and the alternative, by definition of analogous. Detailed visibility to both the analogous programs and the alternative schedules allows one to see the event details which enable the possibilities for increasing schedule time or reducing events. For example, two parallel events could be unraveled into two serial events, hence adding time to the schedule. Risk mitigation strategies are usually performed within the alternative.

16

The DM is typically interested in knowing the trade space between cost, schedule, and performance.   For example, the DM may want to reduce cost and see what the impact is on the technical and schedule risk, and performance. In order to obtain the impact on the schedule risk assessment (if cost is reduced), one needs to know what schedule events and times are affected. The trade-off would have to be clearly defined, and the impact to the schedule would need to be provided by the PM, who developed and knows the details of the schedule (e.g. events, critical path, work breakdown structure).

7. CONCLUSIONS AND PATH FORWARD.

SRDDM has been applied to ongoing AoAs, along with sensitivities for risk mitigation. Some of these include:

- Indirect Fire Protection Capability (IFPC)
- Armored Multi-Purpose Vehicle (AMPV).

The AMSAA Risk Team is capable of using SRDDM to perform risk mitigation and trade space analysis. However, at the current time AMSAA does not have full visibility and knowledge of the schedule details and events. To overcome this problem, the PM could "team up" with AMSAA to perform risk mitigation and trade space analysis since they have full visibility, knowledge, and control of the schedule.

The AMSAA Risk Team will continue to improve the quality of SRDDM in the future by developing event-driven models (e.g. Work Breakdown Structure, Critical Path, Correlation of events), incorporating technical risk assessment outcomes (time distributions for technologies – Technology Readiness Level, Manufacturing Readiness Level, Integration Readiness Level) and SME input. To execute and develop these event-driven models, AMSAA will work with PMs, SMEs, contractors and any other parties that can add insight into the event-driven process.

The event level approaches depend on two core items:
- Data availability and Visibility to the detailed schedule.
- Credible SME's to evaluate detailed event times and assess correlation, etc.

The phase and event level approaches will be applied to an alternative to determine if there is a difference in results.

THIS PAGE INTENTIONALLY LEFT BLANK.

REFERENCES

1. United States Government Accountability Office, Report to Congressional Committees, Defense Management, "Guidance and Progress Measures Are Needed to Realize Benefits from Changes in DOD's Joint Requirements Process", GAO-12-339, February 2012

2. Law, A.M., "Simulation Modeling and Analysis (Fourth Edition)", McGraw Hill, Boston, 2007

3. Ross, S., "Introduction to Probability Models", New York: Academic Press, 2003

4. United States Patent & Trademark Office, Published Patent # U.S.2011/0054839A1, "Confidence Interval Methodology for Ratio Means (CIM4RM)", March 3, 2011, Inventor: John Nierwinski

5. Nierwinski, J., "Maintainability Data Decision Methodology (MDDM)", AMSAA TR-2011-19, June 2011

6. Broemm, W., Nierwinski, J. and Ellner, P., "Parametric Bootstrap Confidence Intervals on the Reliability of Repairable Systems with Grouped Data", AMSAA SR-2010-27-A, November 2010

7. Efron, B. and Tibshirani, R.J., "An Introduction to the Bootstrap", London U.K.: Chapman & Hall/CRC, 1998

8. Wilson, E. B. Probable Inference, the Law of Succession, and Statistical Inference. J of the American Statistical Assoc. 1927, 22, 209-212.

THIS PAGE INTENTIONALLY LEFT BLANK